DOCUMENT RESUME

ED 209 931 PL 012 646

AUTHOR Gordon, Peter
TITLE Syntactic Acquisition of the Count/Mass Distinction.
INSTITUTION Stanford Univ., Calif. Dept. of Linguistics.
PUB DATE Nov 81
NOTE 10p.; In its Papers and Reports on Child Language Development, Number 20, p70-77, Nov 1981.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Child Language; Children; Error Analysis (Language); *Language Acquisition; Language Research; *Nouns; Plurals; *Semantics; *Syntax
IDENTIFIERS *Count Nouns; *Mass Nouns

ABSTRACT

The count/mass disinction is often considered to be a semantic one because it distinguishes those nouns that refer to countable things from those that refer to non-countable things. However, exceptions indicate that semantic properties alone are not sufficient to determine noun sub-categorization. Therefore, such sub-categorization must be defined syntactically for the mature language user. This syntactic definition is couched in terms of distributional properties such as privileges of occurrence within a set of syntactic contexts. Membership within a category or subcategory is defined over a set of linguistic environments into which lexical insertion of that category is permissible. The acquisition of a set of selection rules on quantifiers is examined that invokes the count/mass distinction and serves to partially define the distinction itself from a syntactic distributional viewpoint. Results of an experiment are presented that suggest certain overgeneralizations over this set of syntactic contexts. The results illustrate that in acquiring the count/mass distinction, children must learn that: (1) mass nouns may not be pluralized; (2) certain quantifiers exclusively select for either count or mass nouns; and (3) certain quantifiers take singular count nouns even though plural quantities are being referred to. (Author/JK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED209931

SYNTACTIC ACQUISITION OF THE COUNT/MASS DISTINCTION

Peter Gordon

from

PAPERS AND REPORTS ON CHILD LANGUAGE DEVELOPMENT

Number Twenty

November 1981

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Stanford Univ.
Dept. of Linguistics
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

FL 012 646

(C) 1981 by The Board of Trustees of the Leland Stanford Junior University

All Rights Reserved

Printed in the United States of America

Syntactic Acquisition of the Count/Mass Distinction *

Peter Gordon

Dept. of Psychology

Massachusetts Institute of Technology

The count/mass distinction is often considered to be a semantic one in that it distinguishes those nouns that refer to countable things such as chairs and tables from non-countable stuff such as water and sand. However, there are clearly exceptions to such a rule as in the case of the superordinate terms such as "furniture", "silverware" and "equipment" which are mass nouns yet refer to a collection of things; similarly in the case of "solution" and "sample" (as in blood-) these are count nouns and yet refer to stuff. If semantic properties alone are not sufficient to determine the sub-categorization of the noun, then such sub-categorization must be defined syntactically for the mature language user.

This syntactic definition will presumably be couched in terms of distributional properties such as privelages of occurence within a set of syntactic contexts (as proposed recently by Maratsos and Chalkley 1981). That is, membership within a category or sub-category is defined over a set of linguistic environments into which lexical insertion of that category is permissible.

To organise lexical items in such a manner requires that the child perform operations such as correlation and comparison over a set of syntactic contexts or "frames". In this paper then, I shall examine the acquisition of a set of selection rules on quantifiers which invoke the count/mass distinction and thus serve to partially define the distinction itself from a syntactic-distributional point of view. In particular, I shall present the results of an experiment which suggest certain overgeneralisations over this set of syntactic contexts.

It is generally accepted that the task of language acquisition is to discover the significant generalisations about the language in the form of a set of rules. When such rules are applied inappropriately to marked or irregular forms, this is taken as evidence that the child has both stored the general rule and is applying it productively (c.f. overgeneralisation of past-tense -ed forms to irregular verbs indicates productive application of this rule). Presumably the conditions under which such overgeneralisations occur depend on such factors as: (i) Regularity, (ii) Token-frequency of regular forms in the input data, (iii) Proportion of regular to irregular forms, and (iv) "Generality" which may be defined as the diversity of contexts allowable by the regular rule. That is, a more general rule may be more prone to overgeneralisation than a restrictive one.

In examining the selection rules for quantifiers, there are two main features specified: 1) the marking of the noun as +COUNT or

+MASS, and 2) The presence or absence of the plural. With regard to these two features, quantifiers fall into four main groups as shown in Fig. 1.

Fig 1. Quantifier Selection Rules.

REGULAR			IRREGULAR	
QUANTIFIER	SEL. RULE		QUANTIFIER	SEL. RULE
More Less Enough All (the) (1) A lot of Most Some Plenty of Any	$Q \frown N_{[+count]}^{[+pl]}$ $Q \frown N_{[+mass]}^{[+sg]}$	(2)	Few Several Both Many	$Q \frown N_{[+count]}^{[+pl]}$
		(3)	Much	$Q \frown N_{[+mass]}^{[+sg]}$
		(4)	Each Every Either	$Q \frown N_{[+count]}^{[+sg]}$

As may be seen, group (1) quantifiers take plural count nouns and singular mass nouns (e.g., more chairs, more water); group (2) takes plural count nouns only (e.g., several chairs), group (3) takes singular mass nouns only (e.g., much water) and group (4) takes singular count nouns only (e.g., each chair). We may say then that groups (2) to (4) exclusively select for either count or mass nouns whereas group (1) may take both. Thus, by definition, group (1) quantifiers are more "general". Furthermore, group (1) also contains the largest number of quantifiers which have a total token frequency in text of over five times that of any of the other groups (Kucera & Francis 1967). For these reasons, it may be expected that in acquiring these selection rules, children would generalise toward the form of group (1) which is here termed "regular", the other groups being "irregular".

If such generalisation is employed in the acquisition of these rules, one would predict two kinds of errors: Selection Failures and Overpluralisations. Selection failures involve the failure to select exclusively for either count or mass nouns, and thus irregular quantifiers would modify inappropriately sub-categorized nouns, as in: Several Water (group 2); Much Chairs (group 3) and Each Water (group 4).

Overpluralisations would be predicted to occur with group (4) quantifiers: Each, Every and Either for two reasons. Firstly, a generalisation to the form of group (1) rules should lead to such overpluralisation given that regular quantifiers take plural but not singular count nouns; as in more chairs but not *more chair_.

This prediction is naturally constrained by the nature of the child's existing pluralisation rules themselves. Let us hypothesise that the initial form of such rules constitutes a simple mapping of form to function, that is, singular forms mark singular quantities and plural forms mark plural quantities. Whilst being semantically transparent, such a rule will not reflect the language-specific syntactic constraints on pluralisation. Included in such constraints are the fact that for each every and either plural quantities are referred to yet singular forms are used (i.e., "every car" refers to more than one car). Thus there is a second reason to expect that such overpluralisations will occur for these quantifiers (e.g.,*every cars).

Thus, from the hypothesised overgeneralisation, two kinds of errors are predicted: Selection Failures and Overpluralisations. The task designed to elicit such errors was a forced-choice sentence judgement task. Children were given two sentences from a pair of puppets and were told that one would sound okay and the other silly, and that they were to reward the correct one with a gold star. It was hoped that the forced-choice design would be considerably easier for the child than a straight judgement of grammaticality which is often unworkable with younger children. Referents for nouns were miniatures in a doll's house such as: chairs, cars, kettles etc. (count nouns) and soup, corn, coffee, sugar etc. (mass nouns). On a typical item then, the child would be presented with a table and two chairs and would be told that four people were coming to dinner. One of the puppets would then remark: "That's not enough chairs for four people" and the other would say: "That's not enough chair for four people". The child was then to give the gold star to the puppet who said it "okay".

Given the nature of the forced-choice design, it was necessary to construct sentence-pairs which contrasted correct and incorrect quantifier-noun sequences of the appropriate type. These were divided into Regular and Irregular, examples of which are given in fig. 2. The regular pairs involve group (1) quantifiers and contrast the presence or absence of the plural for count and mass nouns (e.g., more chairs vs. more chair and more water vs. more waters). Irregular conditions tested for the two predictions: selection failures and overpluralisation. Selection failures were tested with irregular quantifiers from groups (2) to (4) which modified inappropriately sub-categorized nouns, as in: *many water. These were paired with "semantically equivalent" quantifiers with the appropriate selection features, as in much water. The quantifiers contrasted here were: Much vs Many; Several vs. A lot of; Each vs All the and Every vs All the.

To test for overpluralisations, group (4) quantifiers (each, every and either) were contrasted by the presence or absence of the plural on count nouns (e.g., every chair vs every chairs).

Fig 2, Contrast-Pairs.

		COUNT		MASS
REGULAR		More Chairs		More Water
	vs.	*More Chair	vs.	*More Waters
		OVERPLUR.		SEL. FAILURE
IRREG.		Every Chair		Much Water
	vs.	*Every Chairs	vs.	*Many Water

One further condition was added to the design which was aimed at examining the nature of early pluralisation rules. It was proposed earlier that initial pluralisation rules might constitute a simple "semantic" mapping of form to function. If this is the case, then the child must learn that, in English, mass nouns may not be pluralised (under normal circumstances). If this syntactic rule is not present to constrain pluralisation at an early age, then one would predict that pluralisation would be conditioned by certain object properties such as discreteness and numerosity. That is children would pluralise soley on the basis of certain perceptual properties such as whether what is being refered to is countable or not, and not whether the noun used is count or mass. Thus, half the subjects were given mass noun referents in a discrete and numerous form such as cans of soup (see fig. 3a) whilst the other half were given referents of a more "continuous", "stuff-like" form such as soup, loose in a pot (fig. 3b). Errors here would be predicted to occur in regular-mass contrasts in which the incorrect frame involves the pluralisation of the mass noun (i.e., *some soups).

Fig 3a, "Soup" -Discrete Referent (DM)

Fig 3b, "Soup" -Continuous Referent (CM)

We thus have three predictions. Within the contrasts using irregular quantifiers, selection failures and overpluralisations are predicted. With the regular-mass contrasts, overpluralisation of mass nouns are predicted when referents are discrete and numerous but not when in a continuous form.

A total of 48 children were tested with ages ranging from 3;3 to 5;10. Each child received 12 sentence-pairs plus training. These were balanced for Regular/Irregular (quantifiers) and Count/Mass (nouns). Half of the subjects were assigned to the discrete-mass condition and the other half to the continuous-mass condition, hereafter refered to as: DM and CM respectively. Overall results comparing regular and irregular contrasts for count and mass nouns are shown in fig 4. For count nouns there is a significant difference ($p = .002$) with greater errors in the irregular conditions (i.e., overpluralisations and selection failures). For mass nouns, it is necessary to partial out errors in the regular conditions attributable to the discrete vs. continuous dimension with respect to referents. That is there is a conflation of the two types of errors being tested. This partialisation is shown in fig 5, in which it can be seen that, in the CM conditions (left columns) there are more errors in the irregular conditions ($p = .037$), but in the DM conditions there is no significant difference, presumably reflecting errors due to the CM vs DM contrast which will be dealt with later.

Fig 4, Error Rates in Regular and Irregular Conditions.

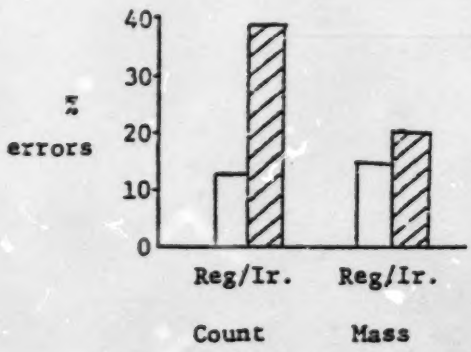
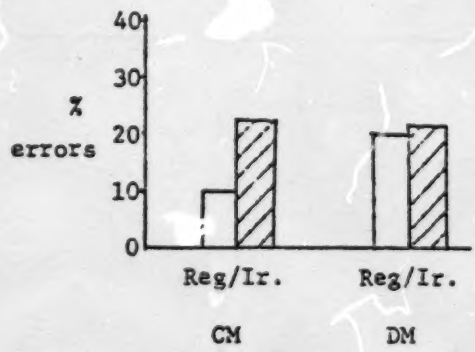


Fig 5, Error Rates Among CM & DM Groups in Mass Noun Conditions.



Within the irregular contrasts, there are two kinds of errors: selection failures and overpluralisations. These are distinguished in fig 6, and fig 7 provides an age-trend analysis. Notice that errors do not decrease significantly with age which suggests that analysis of these rules continues well into the fifth year and possibly beyond. Also, it can be seen that errors are greater for overpluralisation than for selection failures. This is to be expected in that for overpluralisations, the incorrect sentence conformed to the regular rules

and exemplified a simple semantic mapping for pluralisation (as in every chairs), whereas the correct sentence did not (every chair). One would therefore expect something of a preference for the incorrect form where uncertainty exists. For selection failures, on the other hand, both the correct and the incorrect forms exhibited both of these features (as in much chairs vs. many chairs); therefore where strict selection is not in force, there is no preference for either the correct or the incorrect sentence and there is thus a simple trade-off.

Fig 6, Error Rates: Selection-Failure vs. Overpluralisation.

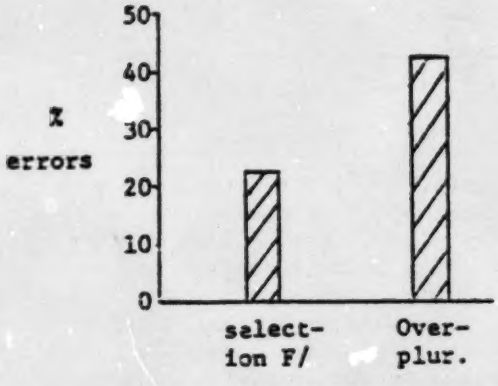
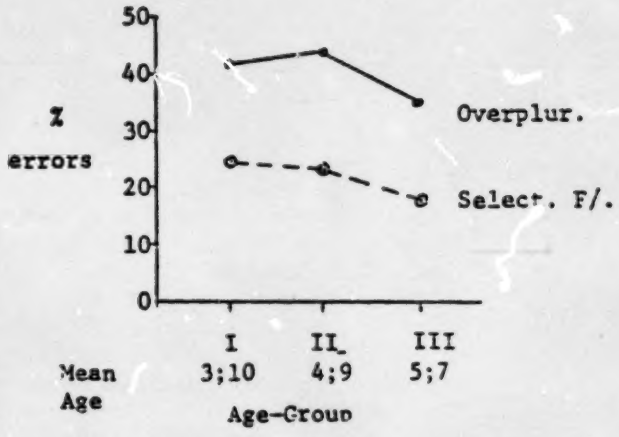
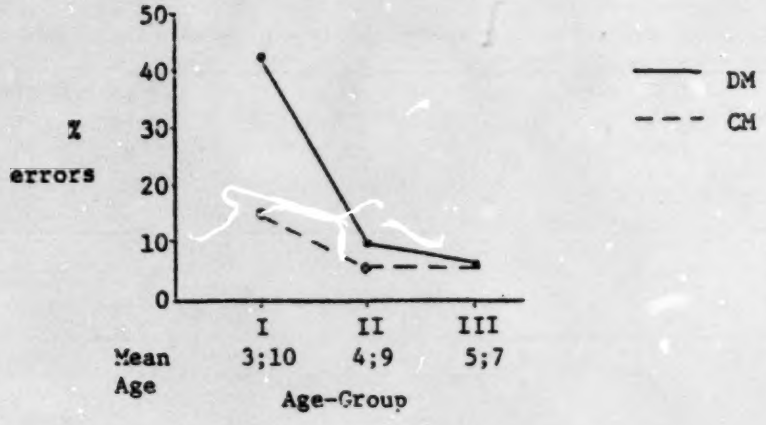


Fig 7, Selection Failures & Overpluralisation Age Trends.



Focussing on the Discrete vs. Continuous contrast for mass nouns, from fig 5 we saw that there were more errors within the DM group than in the CM group, that is, the presence of discrete and numerous mass noun referents contributed to errors of over-pluralisation. This suggests that children were using certain semantic object properties in pluralisation rules and not making reference to the marking of the noun as count or mass. In actual fact, this difference was not overall significant ($p=.08$). However, on examining the age-trends in fig 8, it can be seen that it is mainly the younger children who are making such errors, and for the youngest age group this difference is significant ($p=.04$). In fact the curve shown in fig 8 shows quite a remarkable drop in errors between age groups I and II which suggests something of a "shift" from a semantically based to a syntactically based system of pluralisation. That is, at around age 4½ pluralisation rules start to become constrained by the syntactic sub-categorization of the noun as count or mass.

Fig 8, Regular-Mass Errors by Age in DM and CM Conditions.



To summarize, it would appear that in acquiring these purely formal aspects of the count/mass distinction, the child must learn three things: 1) that mass nouns may not be pluralised, 2) that certain quantifiers exclusively select for either count or mass nouns, and 3) that certain quantifiers take singular count nouns even though plural quantities are being refered to. I have suggested that, based on certain frequency and distributional facts, the child initially employs a generalised rule which is reflected in the errors found. This is thus offered as evidence that acquisition processes and strategies are employed over a set of syntactic contexts which are instrumental in defining the purely syntactic aspects of the count/mass distinction.

The design of this study has deliberately isolated a very narrowly defined aspect of the distinction and has not attempted to address such questions as whether the child uses "semantic bootstrapping" in acquiring the count/mass distinction. Nor is it suggested that these experiments show when the distinction is itself established. In fact, there is evidence from Miller & Ervin-Tripp (unpub) that, distributionally, count and mass nouns are distinguished very early on (c 2yrs) in the use of articles. However, by partialling out the distinct components involved in the count/mass distinction, I hope that it is possible to see how such components interact in acquisition in forming a constellation of rules that serve to define the distinction for the language learner.

* This research was supported by a grant from the Spencer Foundation.I would like to thank Cammy Pane, Carmen Egido and Mara Levin for help with running subjects.Also, I would like to thank the children and staff of the following playgroup/nurseries: YWCA, Boston; Montessorri, Cambridge; Beacon Hill, Boston; Radcliffe, Cambridge; Simmons College, Boston.Finally, I would like to thank Susan Carey for guidance and suggestions.

REFERENCES

Kučera & Francis, (1967) "Computational Analysis of Present Day American English". Brown University Press.

Maratsos, M.P. & Chalkley, (1981) "The Internal Language of Children's Syntax: The Ontogenesis and Representation of Syntactic Categories" In K. Nelson (Ed) Children's Language Vol II. N.Y. Garner Press.

Miller, W. & Ervin-Tripp, S. (unpub) Manuscript report on longitudinal study of language development circa 1963, University of California, Berkeley.